Quelques propriétés d'un estimateur à noyau du quantile conditionnel pour des données associées

Some properties of a kernel conditional quantile estimator for associated data

Latifa Adjoudj[1], Abdelkader Tatachak[1]

[1]Laboratory MSTD, USTHB, Algiers, Algeria, ladjoudj@usthb.dz, atatachak@usthb.dz

RÉSUMÉ. Le présent article vise à établir certaines propriétés asymptotiques d'estimateurs à noyau de la fonction de distribution conditionnelle et du quantile conditionnel lorsque les observations de durée de vie et les covariables sont associées.

ABSTRACT. This paper aims to establish some asymptotic properties of kernel estimators of the conditional distribution function and the conditional quantile when the lifetime observations and the covariates are associated.

MOTS-CLÉS. Association, estimateur à noyau, consistance uniforme forte, quantile conditionnel.

KEYWORDS. Association, conditional quantile, kernel estimator, strong uniform consistency.

Introduction

The study of conditional quantile functions is appealing for a number of reasons, particularly because it provides a more comprehensive picture of the conditional distribution of a dependent variable than traditional regression that restricts attention to the conditional mean function only. It is well known that the quantile function can give a good description of the data, see [CHAUDHURI et al. 1997]. An important application of conditional quantiles is that they provide reference curves or surfaces and conditional prediction intervals that are widely used in many different areas, for example :

— Medicine : reference growth curves for children's height and weight as a function of age ;
— Economics : reference curves to study discrimination effects and trends in income inequality ;
— Ecology : to observe how some covariates can affect limiting sustainable population size ;
— Lifetime analysis : to assess influence of risk factors on survival curves.

A vast literature is dedicated to the nonparametric estimation of conditional quantile.To quote only a few of them, we recall that [MEHRA et al. 1991] and [XIANG 1996] gave the almost sure (a.s.) convergence of a kernel type conditional quantile estimator and its asymptotic normality. [HONDA 2000] dealt with $\alpha-$mixing processes and proved the uniform convergence and asymptotic normality using the local polynomial fitting method.

Here, we are interested in studying another type of dependence, called association. Recall that real-valued random variables (rv's) $\{Z_i; 1 \leq i \leq N\}$ which are defined on a common probability space $(\Omega, \mathscr{A}, \mathbb{P})$ are said to be associated if for every pair of functions $h_1, h_2 : \mathbb{R}^N \rightarrow \mathbb{R}$, which are coordinate-wise non-decreasing and for which $\mathbb{E}\left[h_k^2\left(Z_i, 1 \leq i \leq N\right)\right] < \infty$; $k = 1, 2$, it holds that :

$$cov\left(h_1(Z_i, 1 \leq i \leq N), h_2(Z_j, 1 \leq j \leq N)\right) \geq 0.$$

In classical statistical inference, the rv's of interest are generally assumed to be iid. However, it is more common to have dependent variables in some real life situations. Dependent variables are present in several backgrounds such as medicine, biology and social sciences. The notion of association was firstly introduced by [ESARY et al. 1967] mainly for an application in reliability. For more details on the subject

we refer the reader to the monographs by BULINSKI and SHASHKIN 2007], [OLIVEIRA 2012] and [PRAKASA RAO 2012].

Let $(\mathbf{X}, Y) \in \mathbb{R}^d \times \mathbb{R}$, for which there exist a common unknown joint distribution function (df) $F(\cdot, \cdot)$ and marginals $F_{\mathbf{X}}(\cdot)$ and $F_Y(\cdot)$. In what follow we will denote by $v(\mathbf{x})$, $F(\cdot|\mathbf{x})$ and $f(\cdot|\mathbf{x})$ the probability density function (pdf) of the covariate $\mathbf{X}$, the conditional df and the conditional pdf of Y given $\mathbf{X} = \mathbf{x}$, respectively. The conditional df of Y given $\mathbf{X} = \mathbf{x}$ is $F(y|\mathbf{x}) = \mathbb{E}(\mathbb{1}_{Y \leq y}|\mathbf{X} = \mathbf{x})$ (where $\mathbb{1}_A$ is the indicator function of the set A) which we reformulate as follows :

$$F(y|\mathbf{x}) = \frac{F_1(\mathbf{x}, y)}{v(\mathbf{x})} := \frac{1}{v(\mathbf{x})} \times \frac{\partial F(\mathbf{x}, y)}{\partial \mathbf{x}}. \tag{1}$$

Traditionally, a natural estimator of $F(\cdot|\mathbf{x})$ is the empirical conditional df $F_N(\cdot|\mathbf{x})$, while the estimation of the p-th conditional quantile $q_p(\mathbf{x}) := \inf\{y \in \mathbb{R} : F(y|\mathbf{x}) \geq p\}$, $p \in (0, 1)$ is done via $q_{p,N}(\mathbf{x}) = \inf\{y \in \mathbb{R} : F_N(y|\mathbf{x}) \geq p\}$.

Here, we assume that $\{(\mathbf{X}_i, Y_i); 1 \leq i \leq N\}$ forms a strictly stationary sequence of associated random vectors distributed as $(\mathbf{X}, Y) \in \mathbb{R}^d \times \mathbb{R}$.

The paper is organized as follows. In Section 1, we recall the notations and the definition of our estimators. The main results and the assumptions are listed in Section 2. In Section 3, we evaluate the performance of the estimator on simulated data. The proofs of the results are relegated to Section 4.

1. Notation and estimates

estimation of the conditional df''s is based on the choice of weights. The traditional kernel estimator of $F(y|\mathbf{x})$ is given by $F_N(y|x) = \sum_{i=1}^N w_{i,N}(\mathbf{x}) \mathbb{1}_{Y_i \leq y}$, where $w_{i,N}(.)$ are measurable functions. These weights were introduced by Nadarya-Watson and defined by

$$w_{i,N}(\mathbf{x}) = \frac{K_d\left(\frac{\mathbf{x}-\mathbf{X}_i}{h_{N,K}}\right)}{\sum_{i=1}^N K_d\left(\frac{\mathbf{x}-\mathbf{X}_i}{h_{N,K}}\right)} = \frac{1}{(N h_{N,K}^d)} \frac{K_d\left(\frac{\mathbf{x}-\mathbf{X}_i}{h_{N,K}}\right)}{v_N(\mathbf{x})},$$

with the convention $\frac{0}{0} = 0$, and where

$$v_N(\mathbf{x}) = \frac{1}{N h_{N,K}^d} \sum_{i=1}^N K_d\left(\frac{\mathbf{x} - \mathbf{X}_i}{h_{N,K}}\right), \tag{2}$$

is the classical Parzen-Rosenblatt estimator for v. Here $K_d : \mathbb{R}^d \to \mathbb{R}^+$ is a multivariate kernel such that for any $\mathbf{z}_\ell = (z_\ell^1, z_\ell^2, \ldots, z_\ell^d)^\top \in \mathbb{R}^d$ and a real-valued univariate kernel K

$$\frac{1}{h_{N,K}^d} K_d\left(\frac{\mathbf{z}_\ell}{h_{N,K}}\right) := \prod_{k=1}^d \frac{1}{h_{N,K}} K\left(\frac{z_\ell^k}{h_{N,K}}\right).$$

Furthermore, $\{h_{N,K}\}$ is a sequence of positive constants tending to zero when N tends to infinity, called a bandwidth sequence. Here, $h_{N,K}$ is assumed to be the same regardless of the k-th direction. However, Let us define an estimator for $f(y|\mathbf{x})$ as follows :

$$f_N(y|\mathbf{x}) = \frac{f_N(\mathbf{x}, y)}{v_N(\mathbf{x})} \mathbb{1}_{\{v_N(\mathbf{x}) \neq 0\}},$$

with

$$f_N(\mathbf{x}, y) = \frac{1}{Nh_{N,H}h_{N,K}^d} \sum_{i=1}^{N} K_d\left(\frac{\mathbf{x} - \mathbf{X}_i}{h_{N,K}}\right) H^{(1)}\left(\frac{y - Y_i}{h_{N,H}}\right). \tag{3}$$

Here $H^{(1)}(\cdot)$ denotes a positive kernel function defined on $\mathbb{R}$ and $\int_{-\infty}^{y} H^{(1)}(z)dz =: H(y)$ is a df. The sequences $\{h_{N,K}\} =: h_K$ and $\{h_{N,H}\} =: h_H$ are positive bandwidths that decrease to zero along with N. Now, in view of (1) we have

$$f(\mathbf{x}, y) = \frac{\partial F_1(\mathbf{x}, y)}{\partial y},$$

hence from (3) one may define an estimator for $F_1(\mathbf{x}, y)$ as

$$\begin{aligned} F_{1,N}(\mathbf{x}, y) &= \frac{1}{Nh_H h_K^d} \sum_{i=1}^{N} K_d\left(\frac{\mathbf{x} - \mathbf{X}_i}{h_K}\right) \int_{-\infty}^{y} H^{(1)}\left(\frac{t - Y_i}{h_H}\right) dt \\ &= \frac{1}{nh_K^d} \sum_{i=1}^{N} K_d\left(\frac{\mathbf{x} - \mathbf{X}_i}{h_K}\right) H\left(\frac{y - Y_i}{h_H}\right). \end{aligned} \tag{4}$$

So, the estimators in (2) and (4) enable to derive an estimator for $F(y|\mathbf{x})$, that is

$$F_N(y|\mathbf{x}) = \frac{F_{1,N}(\mathbf{x}, y)}{v_N(\mathbf{x})} \mathbb{1}_{\{v_N(\mathbf{x}) \neq 0\}}. \tag{5}$$

Note that the formulation of the estimator in (5) was introduced by [YU and JONES 1998] as an alternative to the so-called adjusted Nadaraya-Watson estimator of the conditional df. The main motivation was to overcome the crossing problem associated with the kernel weighting estimation when using the indicator function instead of a continuous df.

So, for a fixed $\mathbf{x} \in \mathbb{R}^d$ and $p \in (0, 1)$, a natural estimator (say, $q_{p,N}(\mathbf{x})$) of the p-th conditional quantile can then be defined as

$$q_{p,N}(\mathbf{x}) = \inf\{y \in \mathbb{R} : F_N(y|\mathbf{x}) \geq p\}.$$

2. Main results

Let Ω_0 and $[a, b]$ be two compact subsets of $\Gamma_0 = \{\mathbf{x} \in \mathbb{R}^d; \inf_{\mathbf{x}} v(\mathbf{x}) =: \gamma > 0\}$ and $\mathbb{R}$ respectively. And let us define

$$\eta_{ij} := \sum_{k=1}^{d} \sum_{\ell=1}^{d} cov(X_i^k, X_j^\ell) + 2 \sum_{k=1}^{d} cov(X_i^k, Y_j) + cov(Y_i, Y_j), \tag{6}$$

where X_i^k is the k-th component of $\mathbf{X}_i$; $k = 1, 2, \ldots, d$.

In what follow the letter C will denote a finite positive constant which is allowed to change from line to line.

We will make use of the following assumptions gathered here for easy reference.

2.1. *Assumptions*

(H1) For all $d \geq 1$ and $\beta \in (0,1)$, the bandwidths satisfy
(i) $h_K \to 0$, $Nh_K^{d(1-\beta)+2\beta} \to \infty$ and $\frac{\log^5 N}{Nh_K^d} \to 0$ as $N \to \infty$;
(ii) $h_H \to 0$, $Nh_H h_K^d \to \infty$ as $N \to \infty$;
(H2) The kernel K_d satisfies
(i) K_d is a bounded pdf, compactly supported and Hölder continuous with exponent β;
(ii) $\int_{\mathbb{R}^d} z_k K_d(\mathbf{z})\mathbf{dz} = 0$ for $k = 1, \dots, d$.
(H3) The df $H(\cdot)$ is compactly supported, continuously differentiable and its derivative $H^{(1)}$ is a second-order kernel.
(H4) For all $i \geq 1$, the covariance term in (6) satisfies

$$\sup_{j:|i-j| \geq s} \eta_{ij} =: \rho(s) \leq \tau_0 e^{-\tau s}; \text{ for some positive constants } \tau_0, \tau \text{ and } s.$$

(H5) The pdf $v(\cdot)$ is bounded and satisfies $\sup_{\mathbf{x} \in \Omega_0} \left| \frac{\partial^m v(\mathbf{x})}{\partial x_k \partial x_l^{m-1}} \right| \leq C$ for $k, l = 1, \dots, d$ and $m = 1, 2$.
(H6) For all integers $j \geq 1$, the joint conditional pdf $v_j(\cdot, \cdot)$ of $(\mathbf{X}_1, \mathbf{X}_{1+j})$ exists and satisfies

$$\sup_{(\mathbf{r},\mathbf{s}) \in \Omega_0^2} |v_j(\mathbf{r},\mathbf{s})| \leq C.$$

(H7) The joint pdf $f(\cdot, \cdot)$ of $(\mathbf{X}, Y)$ is bounded and twice continuously differentiable.
(H8) For all integers $j \geq 1$, the joint conditional $f_j(\cdot, \cdot, \cdot, \cdot)$ of $(\mathbf{X}_1, Y_1, \mathbf{X}_{1+j}, Y_{1+j})$ exists and satisfies

$$\sup_{(\mathbf{r},u),(\mathbf{s},t) \in \{\Omega_0 \times [a,b]\}^2} f_j(\mathbf{r}, u, \mathbf{s}, t) \leq C.$$

2.1.1. Comments on the assumptions

(H1) and **(H2)**-**(H3)** are quite usual in kernel estimation setting. Note that the sequences $\{h_K\}$ and $\{h_H\}$ in **(H1)** are not generally equal as it will be seen in the simulations, while several papers suppose their equality to simplify computations. In addition, in view of $nh_H h_K^d \to \infty$ in **(H1)(ii)**, the condition $nh_K^{d(1-\beta)+2\beta} \to \infty$ in **(H1)(i)** becomes superfluous as soon as $d \geq 2$ or for all $d \geq 1$ provided $h_H \leq h_K^\beta$ (β is that in **(H2)(i)**). Thus, **(H1)(i)** is only useful in the case $d = 1$ and $h_H > h_K^\beta$. Assumption **(H4)** quantifies a progressive tendency from dependence to asymptotic independence between "past" and "future". This latter condition was made in Doukhan and Neumann (2007) in order to state an exponential inequality which we apply to prove Proposition 1 hereinafter.
Assumptions **(H5)**-**(H8)** are needed in bias and covariance computation.

2.2. *Strong uniform consistency*

PROPOSITION 1– *Under assumptions (H1)(i)-(ii),(H2),(H3),(H4),(H5) and (H8) we have*

$$\sup_{\mathbf{x} \in \Omega_0} \sup_{a \leq y \leq b} |F_{1,N}(\mathbf{x}, y) - \mathbb{E}[F_{1,N}(\mathbf{x}, y)]| = O\left(\sqrt{\frac{\log N}{Nh_K^d}}\right) \quad \mathbb{P}\text{-a.s., as } N \to \infty.$$

THEOREM 1– *Under assumptions (**H1**)(i)-(ii) and (**H2**)-(**H8**) we have*

$$\sup_{\boldsymbol{x} \in \Omega_0} \sup_{a \leq y \leq b} |F_N(y|\boldsymbol{x}) - F(y|\boldsymbol{x})| = O \left(\sqrt{\frac{\log N}{Nh_K^d}} + h_K^2 + h_H^2 \right) \; \mathbb{P}\text{-a.s., as N} \to \infty,$$

COROLLARY 1– *Under the assumptions of Theorem 1 and for each fixed $p \in (0,1)$ and $\boldsymbol{x} \in \Omega_0$, if the conditional pdf $f(\cdot|\boldsymbol{x})$ and the conditional df $F(\cdot|\boldsymbol{x})$ satisfy $\inf\limits_{x \in \Omega_0} f(q_{p,N}(\boldsymbol{x})|\boldsymbol{x}) > 0$ and $F(a|\boldsymbol{x}) < p < F(b|\boldsymbol{x})$, then we have*

$$\sup_{\boldsymbol{x} \in \Omega_0} |q_{p,N}(\boldsymbol{x}) - q_p(\boldsymbol{x})| = O \left(\sqrt{\frac{\log N}{Nh_K^d}} + h_K^2 + h_H^2 \right) \; \mathbb{P}\text{-a.s., as N} \to \infty.$$

REMARK 1– *The uniform positiveness of the conditional density in Corollary 1 ensures the uniform uniqueness of the conditional quantile, viz.*

$$\forall \varepsilon > 0, \; \exists \beta_0 > 0, \forall \phi_p : \Omega_0 \to \mathbb{R},$$
$$\sup_{\boldsymbol{x} \in \Omega_0} |q_p(\boldsymbol{X}) - \phi_p(\boldsymbol{X})| \geq \varepsilon \Rightarrow \sup_{\boldsymbol{x} \in \Omega_0} |F(q_p(\boldsymbol{x})|\boldsymbol{x}) - F(\phi_p(\boldsymbol{x})|\boldsymbol{x})| \geq \beta_0. \qquad [7]$$

However, Property (7) alone guarantees in part the consistency of the conditional quantile, but does not provide the rate of convergence. The condition on $F(\cdot|\boldsymbol{x})$ means that $a \leq q_{p,N}(\boldsymbol{x}) \leq b$.

3. Experiments with synthetic data

In this section we evaluate the behavior of the studied estimators for some particular conditional quantile functions, considering : the value of p, the sample size for two models (linear and nonlinear). The dimension of the covariate is $d \in \{1, 2\}$.

3.1. *Unidimensional case*

In this case, we study via simulation studies, the performance of the conditional quantile estimator $q_{p,N}(x)\,;\, x \in \mathbb{R}$, for some particular values of p. The measure we used here to quantify the performance is the Global Mean Square Error (GMSE).
For a given functional g and its estimate $\hat{g}_{N,h_K,h_H}$, the GMSE computed along M Monte Carlo trials and a grid of bandwidths h_K and h_H, is defined as

$$\text{GMSE}(h_K, h_H) = \frac{1}{Mm} \sum_{k=1}^{M} \sum_{\ell=1}^{m} (\hat{g}_{N,h_K,h_H,k,\ell}(x) - g(x))^2,$$

where m is a number of equidistant points x_ℓ belonging to a given set and $\hat{g}_{N,h_K,h_H,k,\ell}(x)$ is the value of $\hat{g}_{N,h_K,h_H}(x)$ computed at iteration k with $x = x_\ell\,;\, \ell = 1, 2, \ldots, m$. In order to obtain an associated sequence $(X_i, Y_i)_{i=1,\ldots,N}$, we generate the data according to the following models. Before doing so, we

recall that a rv Z is said to have an inverse Gaussian distribution (also called Wald distribution), with mean $\varpi > 0$ and dispersion parameter $\delta > 0$, denoted by $Z \sim \mathrm{IG}(\varpi, \delta)$ if its pdf is given by

$$f(z) = \left(\frac{\delta}{2\pi z^3} \right)^{\frac{1}{2}} \exp \left\{ -\frac{\delta(z - \varpi)^2}{2\varpi^2 z} \right\} \mathbb{1}_{\{z>0\}}.$$

Note that this distribution is related to the normal distribution via

$$U := \frac{\delta^2 (Z - \varpi/\delta)^2}{Z} \sim \chi^2(1),$$

which means that the rv U is the square of a normally distributed rv.

Then, in order to obtain an observed associated sequence $\{X_i, Y_i; 1 \leq i \leq N\}$, we generate the data according to the following models.

1. Model 1 : *Linear case*

   (a) The covariate X :

      - Generate $(N + 1)$ i.i.d. $IG(1, 1)$ rv's $\{W_i; i = -1, 0, \ldots, N - 1\}$.
      - Given W_i, generate the associated sequence $\{X_i, i = 1, \ldots, N\}$ by $X_i = (W_{i-1} + W_{i-2})/2$.

   (b) The interest variable Y :

      - Generate N i.i.d. $\mathcal{N}(0, 0.01)$ rv's $\{\varepsilon_i; i = 1, \ldots, N\}$.
      - Set $Y_i = 0.5\, X_i + \sigma \varepsilon$.

   Note that under this model, the rv $Z = (Y|X = x) \sim \mathcal{N}(x/2, 0.01)$. For the estimators, we employ the standard Gaussian kernel and $M = 200$. The pairs of bandwidths (h_K, h_H) take their values in the set $[0.05,\ 0.95]$. At the end of the process, we retain the minimal values of the GMSE's computed along the grids as well as the corresponding optimal pairs of bandwidths minimizing the errors.

2. Model 2 : *Non-linear case*

   The same procedure as before (linear case) is followed to measure the performance of the estimators. Note that we use the function $\log$ to preserve the association property. The model is $Y = \log(X) + \varepsilon$. The rv's X and ε follow the same distributions as above whereas $Z = (Y|X = x) \sim \mathcal{N}(\log(x), 0.01)$.

The different values computed in each case are gathered in Table 3.1.

3.2.   *Two-dimensional case (d=2).*

In this subsection, we restrict ourselves to $p = 0.5$, the linear and nonlinear two-dimensional cases are studied.

- Model 1 : *Linear case (d=2)*

		Model 1 : $Y = \frac{1}{2}X + \varepsilon$			Model 2 : $Y = \log(X) + \varepsilon$		
p	N	GMSE	h_H	h_K	GMSE	h_H	h_K
	50	0.0102	0.95	0.2	0.0059	0.95	0.2
0.5	500	0.0016	0.75	0.15	0.0014	0.3	0.15
	50	0.0090	0.15	0.2	0.1972	0.1	0.1
0.25	500	0.0013	0.1	0.1	0.0010	0.05	0.05

Tableau 3.1.: The GMSE's values and the corresponding optimal pairs of bandwidths

1. The covariate X :
   - $X_1(i) = (W_1(i-1) + W_1(i-2))/2$, where $\{W_1(i);\ i = -1, 0, \cdots, N-1\}$ are $N+1$ iid rv's $\sim \mathcal{IG}(1,1)$; (Inverse Gaussian)
   - $X_2(i) = (W_2(i-1) + W_2(i-2))/2$, where $\{W_2(i);\ i = -1, 0, \cdots, N-1\}$ are $N+1$ iid rv's $\sim \mathcal{IG}(1,1)$,

2. The target variable Y :
   - We generate N iid rv's $\{\varepsilon_i;\ i = 1, \cdots, N\}$ with $\varepsilon_i \sim \mathcal{N}(0, 0.01)$.
   - We generate the target variable as follow $Y = 0.5(X_1 + X_2) + \varepsilon$ where $X = (X_1, X_2)^\top$.

- Model 2 : *Non-linear case.* We simulate the covariate as before (linear case) and compute
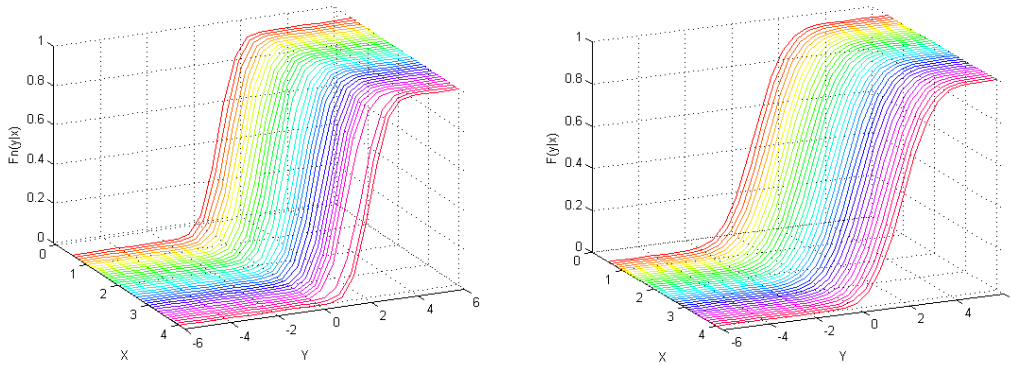$Y(i) = (\log(X_1(i) + 1) + \log(X_2(i) + 1)) + \varepsilon_i \ \{i = 1, \dots, N\}$

Figure 1.: The estimator $F_N(y|x)$ versus $F(y|x)$ for $N = 100$
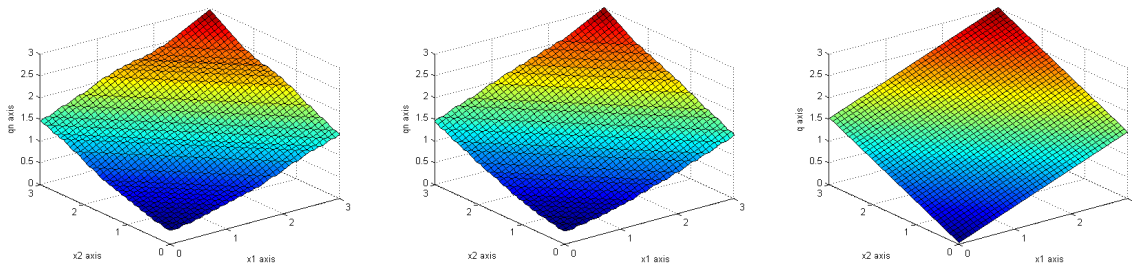

Figure 2.: Model 1 : Conditional median surface versus theoretical linear function (right) for N= 200 and N= 500
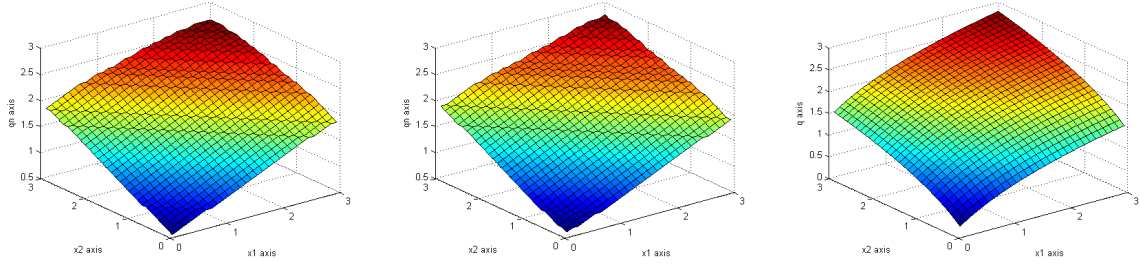
Figure 3.: Model 2 :Conditional median surface versus theoretical nonlinear function (right) for N= 200 and N= 500

4. Auxiliary results and proofs

For notational convenience we set

$$Z_i(\mathbf{x}, y) = K_d\left(\frac{\mathbf{x} - \mathbf{X}_i}{h_K}\right) H\left(\frac{y - Y_i}{h_H}\right) - \mathbb{E}\left[K_d\left(\frac{\mathbf{x} - \mathbf{X}_1}{h_K}\right) H\left(\frac{y - Y_1}{h_H}\right)\right] =: Z_i.$$

Then it is clear that we have

$$F_{1,N}(\mathbf{x}, y) - \mathbb{E}\left[F_{1,N}(\mathbf{x}, y)\right] = \frac{1}{N h_K^d} \sum_{i=1}^{N} Z_i.$$

We recall in the following lemma an exponential inequality stated in Doukhan and Neumann (2007) used in the proof of proposition 1.

LEMMA 1– *Under assumptions (H2)(i),(H4) and (H8), there exist constants M_0, L_1, L_2 and $\mu, \lambda \geq 0$ such that for all $(s_1, \ldots, s_u) \in \mathbb{N}^u$ and all $(t_1, \ldots, t_v) \in \mathbb{N}^v$ with $1 \leq s_1 \leq \ldots \leq s_u \leq t_1 \leq \ldots \leq t_v \leq n$, we have*

(a) $\mathrm{cov}\left(\prod_{i=s_1}^{s_u} Z_i, \prod_{j=t_1}^{t_v} Z_j\right) \leq C^{u+v} h_K^d h_H^{\frac{2}{d+1}} uv \left(\rho(t_1 - s_u)\right)^{\frac{d}{2d+2}}$;

(b) $\sum_{s=0}^{\infty} (s+1)^k (\rho(s))^{\frac{d}{2d+2}} \leq L_1 L_2^k (k!)^{\mu}, \forall k \geq 0$;

(c) $\mathbb{E}\left(|\psi_i|^k\right) \leq (k!)^{\lambda} M_0^k, \forall k \geq 0.$

4.1. *Proof of Lemma 1*

To prove this lemma we first recall that $\Phi_m : \mathbb{R}^m \to \mathbb{R}$ is said to be a Lipschitz function if

$$Lip(\Phi_m) = \sup_{\mathbf{x} \neq \mathbf{y}} \frac{|\Phi_m(\mathbf{x}) - \Phi_m(\mathbf{y})|}{\|\mathbf{x} - \mathbf{y}\|_1} \leq C,$$

where $\|\mathbf{z}\|_1 = |z_1| + |z_2| + \cdots + |z_m|$. For such a function, the partial Lipschitz constants (see Bulinski and Shashkin 2007, Definition 5.1, p. 88) are

$$Lip_i(\Phi_m) = \sup_{\substack{z_1,\ldots,z_m,z_i'\in\mathbb{R} \\ z_i\neq z_i'}} \frac{|\Phi_m(z_1,\ldots,z_{i-1},z_i,z_{i+1},\ldots,z_m) - \Phi_m(z_1,\ldots,z_{i-1},z_i',z_{i+1},\ldots,z_m)|}{|z_i - z_i'|}.$$

[8]

So, to state item (a) we set

$$\Phi_u := \prod_{i=s_1}^{s_u} Z_i \;\; \text{and} \;\; \Phi_v := \prod_{j=t_1}^{t_v} Z_j.$$

Firstly by using Theorem 5.3 in [BULINSKI and SHASHKIN 2007], the definition in (6) and the fact that K_d and H are Lipschitz functions we may write

$$cov\,(\Phi_u, \Phi_v) \leq \sum_{i=s_1}^{s_u}\sum_{j=t_1}^{t_v} Lip_i(\Phi_u)Lip_j(\Phi_v)\eta_{ij}.$$

Next, in view of the definition in (8) and by a simple computation but long enough, we obtain

$$Lip_i(\Phi_u) \leq \frac{C}{h_K h_H}\left(2\,\|K_d\|_\infty\right)^{u-1}$$

and

$$Lip_i(\Phi_v) \leq \frac{C}{h_K h_H}\left(2\,\|K_d\|_\infty\right)^{v-1}$$

with

$$C = \max\left\{h_H\|K\|_\infty^{d-1}Lip(K), h_K\|K_d\|_\infty Lip(H)\right\}.$$

where $\|\cdot\|_\infty$ stands for the sup norm. Hence by Assumption (**H**4) and the stationarity we get

$$cov\,(\Phi_u, \Phi_v) \;\leq\; \frac{C^2}{(h_K h_H)^2}\|K_d\|_\infty^{u+v-2}\sum_{i=s_1}^{s_u}\sum_{j=t_1}^{t_v}\eta_{ij}$$

$$\leq\; \frac{C}{(h_K h_H)^2}\|K_d\|_\infty^{u+v}uv\rho(t_1 - s_u) =: cov_1.$$

[9]

On the other hand, under (**H**8), a change of variable and (**H**1)(i) we have

$$|cov\,(Z_i, Z_j)| \;=\; |\mathbb{E}(Z_i Z_j)|$$

$$\leq\; \left(\int_{\mathbb{R}^d}\int_{\mathbb{R}}\int_{\mathbb{R}^d}\int_{\mathbb{R}} K_d\left(\frac{\mathbf{x}-\mathbf{u}}{h_K}\right) H\left(\frac{y-r}{h_H}\right) K_d\left(\frac{\mathbf{x}-\mathbf{s}}{h_K}\right) H\left(\frac{y-t}{h_H}\right)\right.$$

$$\times f_j(\mathbf{u}, r, \mathbf{s}, t)\mathbf{du}dr\mathbf{ds}dt\Big) + \left(\int_{\mathbb{R}^d}\int_{\mathbb{R}} K_d\left(\frac{\mathbf{x}-\mathbf{u}}{h_K}\right) H\left(\frac{y-r}{h_H}\right) f(\mathbf{u}, r)\mathbf{du}dr\right)$$

$$\times \int\limits_{\mathbb{R}^d} \int\limits_{\mathbb{R}} K_d \left(\frac{\mathbf{x} - \mathbf{s}}{h_K}\right) H \left(\frac{y - t}{h_H}\right) f(\mathbf{s}, t) \mathbf{ds} dt \Bigg)$$
$$= O \left(h_K^{2d} h_H^2\right),$$

which helps to obtain

$$|cov\,(\Phi_u, \Phi_v)| \leq C^{u+v} h_K^{2d} h_H^2 =: cov_2. \qquad [10]$$

Then by combining (9) and (10), we get

$$|cov\,(\Phi_u, \Phi_v)| \leq C^{u+v} h_K^d h_H^{\frac{2}{d+1}} uv \left(\rho(t_1 - s_u)\right)^{\frac{d}{2d+2}}. \qquad [11]$$

For items (b) and (c), the proofs are similar to those of Proposition 8 in [DOUKHAN and NEUMANN 2007]. It suffices to choose $\mu = 1$, $L_1 = L_2 = (1 - e^{\frac{-\tau d}{2d+2}})^{-1}$, $\lambda = 0$ and use Assumption (**H**4). Then we omit them. $\qquad\square$

4.2. *Proof of Proposition 1*

The proof is based on the following observation : The compacts sets Ω_0 and $[a, b]$ can be covered respectively by a finite number $d_{x,N}$ of balls B_k centred at $\mathbf{x}_k \in \mathbb{R}^d$ such that

$$\max_{1 \leq k \leq d_{x,N}} \|\mathbf{x} - \mathbf{x}_k\| \leq \left(N^{-\frac{1}{2}} h_K^{\frac{d}{2} + \beta}\right)^{\frac{1}{\beta}} =: \omega_N^d$$

and $d_{y,n}$ intervals J_ℓ centred at y_ℓ satisfying

$$\max_{1 \leq \ell \leq d_{y,N}} |y - y_\ell| \leq h_H (n h_K^d)^{-\frac{1}{2}} =: \zeta_n.$$

Since Ω_0 and $[a, b]$ are bounded, one can find two positive constants M_x and M_y such that $d_{x,N} \omega_n^d \leq M_x$, $d_{y,N} \zeta_N \leq M_y$ and $M_x M_y \leq C$.
Hence we consider the following decomposition

$$\sup_{\mathbf{x} \in \Omega_0} \sup_{a \leq y \leq b} |F_{1,N}(\mathbf{x}, y) - \mathbb{E}[F_{1,N}(\mathbf{x}, y)]| \leq \max_{1 \leq k \leq d_{x,N}} \sup_{\mathbf{x} \in B_k} \sup_{a \leq y \leq b} |F_{1,N}(\mathbf{x}, y) - F_{1,N}(\mathbf{x}_k, y)|$$
$$+ \max_{1 \leq k \leq d_{x,N}} \sup_{\mathbf{x} \in B_k} \sup_{a \leq y \leq b} |\mathbb{E}[F_{1,N}(\mathbf{x}, y)] - \mathbb{E}[F_{1,N}(\mathbf{x}_k, y)]|$$
$$+ \max_{1 \leq k \leq d_{x,N}} \max_{1 \leq \ell \leq d_{y,N}} \sup_{y \in J_\ell} |F_{1,N}(\mathbf{x}_k, y_\ell) - F_{1,N}(\mathbf{x}_k, y)|$$
$$+ \max_{1 \leq k \leq d_{x,N}} \max_{1 \leq \ell \leq d_{y,N}} \sup_{y \in J_\ell} |\mathbb{E}[F_{1,N}(\mathbf{x}_k, y_\ell)] - \mathbb{E}[F_{1,N}(\mathbf{x}_k, y)]|$$
$$+ \max_{1 \leq k \leq d_{x,N}} \max_{1 \leq \ell \leq d_{y,N}} |F_{1,N}(\mathbf{x}_k, y_\ell) - \mathbb{E}[F_{1,N}(\mathbf{x}_k, y_\ell)]|$$
$$= \mathscr{I}_{1N} + \mathscr{I}_{1N}' + \mathscr{I}_{2N} + \mathscr{I}_{2N}' + \mathscr{I}_{3N}. \qquad [12]$$

Under assumptions (**H**1)(i) we have

$$|F_{1,N}(\mathbf{x}, y) - F_{1,N}(\mathbf{x}_k, y)| \leq C \frac{\|\mathbf{x} - \mathbf{x}_k\|^\beta}{h_K^\beta} \frac{1}{N h_K^d} \sum_{i=1}^N H \left(\frac{y - Y_i}{h_H}\right)$$

$$\le C \frac{(\omega_N^d)^\beta}{h_K^{d+\beta}}$$

$$= O\left(\left(Nh_K^d\right)^{-1/2}\right). \tag{13}$$

Similar arguments as above lead to the same bound for $\mathscr{I}'_{1N}$.

Under (**H2**), and according to Markov inequality and the Dominated Convergence Theorem we obtain the same upper-bound for $\mathscr{I}_{2N}$. We have

$$
\begin{aligned}
|F_{1,n}(\mathbf{x}_k, y_\ell) - F_{1,N}(\mathbf{x}_k, y)| \quad &\le C \frac{|y - y_\ell|}{h_H} \frac{1}{Nh_K^d} \sum_{i=1}^{N} K_d\left(\frac{\mathbf{x}_k - \mathbf{X}_i}{h_K}\right) \\
&\le C \frac{\zeta_N}{h_H} \frac{1}{Nh_K^d} \sum_{i=1}^{N} K_d\left(\frac{\mathbf{x}_k - \mathbf{X}_i}{h_K}\right) \\
&\le C \frac{\zeta_N}{h_H)} v_N(\mathbf{x}) \\
&= O\left(\left(Nh_K^d\right)^{-1/2}\right).
\end{aligned}
\tag{14}
$$

The term $\mathscr{I}'_{2N}$ is upper-bounded similarly as $\mathscr{I}_{2N}$. We now turn to the term $\mathscr{I}_{3n}$ we use Theorem 1 of [DOUKHAN and NEUMANN 2007]. For this, consider the centered random functions $Z_i(\mathbf{x}_k, y_\ell); i = 1, \dots, n$ satisfying the items in Lemma 1. Then for any $\varepsilon > 0$ we have

$$
\mathbb{P}\left(\sum_{i=1}^{N} Z_i(\mathbf{x}_k, y_\ell) \ge \varepsilon\right) \le \exp\left(-\frac{\varepsilon^2/2}{A_N + B_N^{1/(\mu+\lambda+2)} \varepsilon^{(2\mu+2\lambda+3)/(\mu+\lambda+2)}}\right), \tag{15}
$$

where A_N and B_N can be chosen such that

$$
A_N \le \sigma_N^2 \text{ with } \sigma_N^2 := \mathrm{var}\left(\sum_{i=1}^{N} Z_i(\mathbf{x}_k, y_\ell)\right) \text{ and } B_N = 2CL_2\left(\frac{2^{4+\mu+\lambda} CN h_K^d h_H^{\frac{2}{d+1}} L_1}{A_N} \vee 1\right).
$$

For this purpose, we first evaluate σ_N^2.

$$
\begin{aligned}
\sigma_N^2 \quad &= \quad \mathrm{var}\left(\sum_{i=1}^{N} Z_i(\mathbf{x}_k, y_\ell)\right) \\
&= \quad \mathrm{var}\left(Nh_K^d \tilde{F}_{1,n}(\mathbf{x}_k, y_\ell)\right) \\
&= \quad N\,\mathrm{var}\left(K_d\left(\frac{\mathbf{x}_k - \mathbf{X}_1}{h_K}\right) H\left(\frac{y_\ell - Y_1}{h_H}\right)\right) \\
&\quad + \sum_{i=1}^{N}\sum_{j=1}^{N} cov\left(K_d\left(\frac{\mathbf{x}_k - \mathbf{X}_i}{h_K}\right) H\left(\frac{y_\ell - Y_i}{h_H}\right), K_d\left(\frac{\mathbf{x}_k - \mathbf{X}_j}{h_K}\right) H\left(\frac{y_\ell - Y_j}{h_H}\right)\right) \\
&=: \quad \mathscr{V} + \mathscr{S}.
\end{aligned}
$$

On the one hand we have

$$
\mathscr{V} \quad = \quad N\left\{\mathbb{E}\left[K_d^2\left(\frac{\mathbf{x}_k - \mathbf{X}_1}{h_K}\right) H^2\left(\frac{y_\ell - Y_1}{h_H}\right)\right] - \mathbb{E}^2\left[K_d\left(\frac{\mathbf{x}_k - \mathbf{X}_1}{h_K}\right) H\left(\frac{y_\ell - Y_1}{h_H}\right)\right]\right\}
$$

$$=: N\left\{\mathscr{V}_1 - \mathscr{V}_2\right\}.$$

By a change of variable, a Taylor expansion, assumptions (**H1**)(i) (ii),(**H2**)(i) and (**H7**) we get

$$
\begin{aligned}
\mathscr{V}_1 &= \mathbb{E}\left[K_d^2\left(\frac{\mathbf{x}_k - \mathbf{X}_1}{h_K}\right)H^2\left(\frac{y_\ell - Y_1}{h_H}\right)\right] \\
&\leq \int_{\mathbb{R}^d} K_d^2\left(\frac{\mathbf{x}_k - \mathbf{u}}{h_K}\right)\left(\int_{\mathbb{R}} f(\mathbf{u}, s)ds\right)\mathbf{du}, \quad \text{car } H(\cdot) \\
&\leq C h_K^d \int_{\mathbb{R}^d} K_d^2(\mathbf{z})v(\mathbf{x}_k - \mathbf{z}h_K)\mathbf{dz} \\
&= O(h_K^d).
\end{aligned}
$$

For $\mathscr{V}_2$ we work as for $\mathscr{V}_1$. Under assumptions (**H1**)(i) (ii),(**H2**)(i),(**H7**), a change of variable and a Taylor expansion, we get

$$
\begin{aligned}
\mathscr{V}_2 &= \mathbb{E}^2\left[K_d\left(\frac{\mathbf{x}_k - \mathbf{X}_1}{h_K}\right)H\left(\frac{y_\ell - Y_1}{h_H}\right)\right] \\
&= O(h_K^{2d}),
\end{aligned}
$$

Thus $\mathscr{V} = O(Nh_K^d)$.

Now to deal with $\mathscr{S}$, let us define $\mathscr{B}_1 = \{(i,j); 1 \leq |i-j| \leq w_N\}$ and $\mathscr{B}_2 = \{(i,j); w_N + 1 \leq |i-j| \leq N - 1\}$, where $w_N = o(N)$. Then

$$
\begin{aligned}
\mathscr{S} &= \sum_{i=1}^{N}\sum_{j\in\mathscr{B}_1} cov\left(K_d\left(\frac{\mathbf{x}_k - \mathbf{X}_i}{h_K}\right)H\left(\frac{y_\ell - Y_i}{h_H}\right), K_d\left(\frac{\mathbf{x}_k - \mathbf{X}_j}{h_K}\right)H\left(\frac{y_\ell - Y_j}{h_H}\right)\right) \\
&\quad + \sum_{i=1}^{N}\sum_{j\in\mathscr{B}_2} cov\left(K_d\left(\frac{\mathbf{x}_k - \mathbf{X}_i}{h_K}\right)H\left(\frac{y_\ell - Y_i}{h_H}\right), K_d\left(\frac{\mathbf{x}_k - \mathbf{X}_j}{h_K}\right)H\left(\frac{y_\ell - Y_j}{h_H}\right)\right) \\
&=: \mathscr{S}_1 + \mathscr{S}_2.
\end{aligned}
$$

From (10) we get

$$\mathscr{S}_1 = O\left(Nw_N h_K^{2d} h_H^2\right).$$

Then under Assumption (**M**) and taking $u = v = 1$ in (11), we can write

$$
\begin{aligned}
\mathscr{S}_2 &\leq \sum_{i=1}^{N}\sum_{j\in\mathscr{B}_2} C^2 h_K^d h_H^{\frac{2}{d+1}}\left(\rho(|i-j|)\right)^{\frac{d}{2d+2}} \\
&\leq CN h_K^d h_H^{\frac{2}{d+1}} \sum_{l\in\mathscr{B}_2} e^{-\frac{\tau|i-j|d}{2d+2}} \\
&\leq CN h_K^d h_H^{\frac{2}{d+1}} \int_{w_N}^{N} e^{-\frac{\tau d\xi}{2d+2}}d\xi
\end{aligned}
$$

$$= O\left(Nh_K^d h_H^{\frac{2}{d+1}} e^{-\frac{\tau d w_N}{2d+2}}\right).$$

So, under Assumption **(B)**(i)(ii) and choosing $w_N = O\left(h_K^{\nu_1-d}h_H^{\nu_2-1}\right)$ with $0 < \nu_1 < d$ and $0 < \nu_2 < 1$, the terms $\mathscr{S}_1$ and $\mathscr{S}_2$ become of order $o(Nh_K^d h_H)$ and $o(Nh_K^d h_H^{\frac{2}{d+1}})$, respectively. Hence, $\mathscr{S} = o(Nh_K^d h_H^{\frac{2}{d+1}})$ and $\sigma_N^2 = O(Nh_K^d)$ and therefore we can take $A_N = O(Nh_K^d)$. In addition, by choosing μ, λ, L_1 and L_2 as before in proving Lemma 1, we get $B_N = O(1)$. Consequently (15) becomes

$$\mathbb{P}\left(\sum_{i=1}^{N} Z_i(\mathbf{x}_k, y_\ell) \geq \varepsilon\right) \leq \exp\left(-\frac{\varepsilon^2/2}{CNh_K^d + \varepsilon^{\frac{5}{3}}}\right). \tag{16}$$

Furthermore, if we choose $\varepsilon = \varepsilon_0 \sqrt{\frac{\log N}{Nh_K^d}}$; $\varepsilon_0 > 0$, then from (16) we have

$$
\begin{aligned}
\mathbb{P}\left(\mathscr{I}_{3N} \geq \varepsilon\right) &= \mathbb{P}\left(\max_{1 \leq k \leq d_{x,N}} \max_{1 \leq \ell \leq d_{y,N}} \left|\sum_{i=1}^{N} Z_i(\mathbf{x}_k, y_\ell)\right| \geq \varepsilon Nh_K^d\right) \\
&\leq 2d_{x,N}d_{y,N} \exp\left(-\frac{(\varepsilon Nh_K^d)^2/2}{CNh_K^d + (\varepsilon Nh_K^d)^{\frac{5}{3}}}\right) \\
&\leq C(\omega_N^d)^{-1}\zeta_N^{-1} \exp\left(-\frac{(\varepsilon_0^2 \log N)/2}{C + \varepsilon_0^{\frac{5}{3}}\left(\frac{\log^5 N}{Nh_K^d}\right)^{1/6}}\right) \\
&\leq C\left(N^{-\frac{1}{2}}h_K^{\frac{d}{2}+\beta}\right)^{\frac{-1}{\beta}}\left((Nh_K)^{-\frac{1}{2}}h_H\right)^{-1} N^{-C\varepsilon_0^2} \\
&= \left(Nh_K^{d(1-\beta)+2\beta}\right)^{-\frac{1}{2\beta}}(Nh_H)^{-1} O\left(N^{-C\varepsilon_0^2+\frac{1}{\beta}+\frac{3}{2}}\right).
\end{aligned} \tag{17}
$$

Then, for an appropriate choice of ε_0 and Assumption(**H**1) (i) (ii), the term in (17) becomes an $O(N^{-\theta})$ with $\theta > 1$. So, we obtain

$$\mathbb{P}\left(\mathscr{I}_{3N} \geq \varepsilon_0 \sqrt{\frac{\log N}{Nh_K^d}}\right) = O(n^{-\theta}). \tag{18}$$

Hence, according to (12),(13),(14) and (18), if we set $\varepsilon_1 = 5\varepsilon_0$, it follows that

$$\sum_{n \geq 1} \mathbb{P}\left(\sup_{\mathbf{x} \in \Omega_0} \sup_{a \leq y \leq b} |F_{1,N}(\mathbf{x}, y) - \mathbf{E}[F_{1,N}(\mathbf{x}, y)]| \geq \varepsilon_1 \sqrt{\frac{\log N}{Nh_K^d}}\right) \leq \sum_{N \geq 1} \mathbb{P}\left(\mathscr{I}_{3N} \geq \varepsilon_0 \sqrt{\frac{\log N}{Nh_K^d}}\right) < \infty.$$

Thus by the Borel-Cantelli lemma, we conclude that

$$\sup_{\mathbf{x} \in \Omega_0} \sup_{a \leq y \leq b} |F_{1,N}(\mathbf{x}, y) - \mathbb{E}\left[F_{1,N}(\mathbf{x}, y)\right]| = O\left(\sqrt{\frac{\log N}{Nh_K^d}}\right) \quad \mathbb{P}\text{-a.s., as } N \to \infty.$$

The proof of Proposition 1 is complete. $\qquad\square$

LEMMA 2– *Under assumptions (H2),(H3) and (H7) we have*

$$\sup_{x\in\Omega_0}\sup_{a\leq y\leq b}|\mathbb{E}[F_{1,N}(\boldsymbol{x},y)]-F_1(\boldsymbol{x},y)|=O(h_K^2+h_H^2).$$

4.3. *Proof of Lemma 2*

The proof is based on integration by parts, a change of variable and a Taylor expansion, indeed we have

$$
\begin{aligned}
\mathbb{E}[F_{1,n}(\mathbf{x},y)]-F_1(\mathbf{x},y) &= \mathbb{E}\left[\frac{1}{nh_K^d}\sum_{i=1}^N K_d\left(\frac{\mathbf{x}-\mathbf{X}_i}{h_K}\right)H\left(\frac{y-Y_i}{h_H}\right)\right]-F_1(\mathbf{x},y)\\
&= \frac{1}{h_K^d}\int_{\mathbb{R}^d} K_d\left(\frac{\mathbf{x}-\mathbf{r}}{h_K}\right)\left(\int_{\mathbb{R}} H^{(1)}(s)F_1(\mathbf{r},y-sh_H)ds\right)\mathbf{dr}-F_1(\mathbf{x},y)\\
&= \int_{\mathbb{R}^d}\int_{\mathbb{R}} K_d(\mathbf{u})H^{(1)}(s)\left(F_1(\mathbf{x}-\mathbf{u}h_K,y-sh_H)-F_1(\mathbf{x},y)\right)\mathbf{du}ds.
\end{aligned}
$$

Then under assumptions (**H2**),(**H3**) and (**H7**) we get

$$\sup_{\mathbf{x}\in\Omega_0}\sup_{a\leq y\leq b}|\mathbb{E}[F_{1,n}(\mathbf{x},y)]-F_1(\mathbf{x},y)|=O(h_K^2+h_H^2).$$

The proof of Lemma 2 is finished. □

LEMMA 3– *Under assumptions (H1)(i),(H2)and (H4)−(H6) we have*

$$\sup_{x\in\Omega_0}|v_N(\boldsymbol{x})-v(\boldsymbol{x})|=O\left(\sqrt{\frac{\log N}{Nh_K^d}}+h_K^2\right)\ \mathbb{P}\text{-a.s., as N}\to\infty.$$

4.4. *Proof of Lemma 3*

We have

$$\sup_{\mathbf{x}\in\Omega_0}|v_N(\mathbf{x})-v(\mathbf{x})|\ \leq\ \sup_{\mathbf{x}\in\Omega_0}|v_N(\mathbf{x})-\mathbb{E}[v_N(\mathbf{x})]|+\sup_{\mathbf{x}\in\Omega_0}|\mathbb{E}[v_N(\mathbf{x})]-v(\mathbf{x})|$$

suggests to treat the first term in the right-hand side by following the same steps and similar arguments used in proving Proposition 1. Hence we get

$$\sup_{\mathbf{x}\in\Omega_0}|v_N(\mathbf{x})-\mathbb{E}[v_N(\mathbf{x})]|=O\left(\sqrt{\frac{\log N}{Nh_K^d}}\right)\ \mathbb{P}\text{-p.s., as N}\to\infty.$$

Furthermore, under assumptions (**H1**)(i),(**H2**)and (**H5**) and using a Taylor expansion, we get

$$\sup_{\mathbf{x}\in\Omega_0}|\mathbb{E}[v_N(\mathbf{x})]-v(\mathbf{x})|=O(h_K^2).$$

By combining 19 and 19 we end the proof of Lemma 3. □

4.5. *Proof of Theorem 1*

The proof is based on the following decomposition.

$$
\sup_{x \in \Omega_0} \sup_{a \leq y \leq b} |F_N(y|\mathbf{x}) - F(y|\mathbf{x})| \leq \frac{1}{\gamma - \sup_{\mathbf{x} \in \Omega_0} |v_N(\mathbf{x}) - v(\mathbf{x})|} \left\{ \sup_{\mathbf{x} \in \Omega_0} \sup_{a \leq y \leq b} |F_{1,N}(\mathbf{x}, y) - \mathbb{E}[F_{1,N}(\mathbf{x}, y)]| \right.
$$
$$
+ \sup_{\mathbf{x} \in \Omega_0} \sup_{a \leq y \leq b} |\mathbb{E}[F_{1,N}(\mathbf{x}, y)] - F_{1,N}(\mathbf{x}, y)|
$$
$$
\left. + \gamma^{-1} \sup_{\mathbf{x} \in \Omega_0} \sup_{a \leq y \leq b} F_1(\mathbf{x}, y) \times \sup_{x \in \Omega_0} |v_N(\mathbf{x}) - v(\mathbf{x})| \right\}
$$

Then by using Proposition 1, Lemma 2 and Lemma 3 we get the result.

4.6. *Proof of Corollary 1*

A Taylor series expansion of $F(q_{p,N}(\cdot)|\cdot)$ around $q_p(\cdot)$ gives

$$
F(q_{p,N}(\mathbf{x})|\mathbf{x}) - F(q_p(\mathbf{x})|\mathbf{x}) = (q_{p,N}(\mathbf{x}) - q_p(\mathbf{x})) f\left(q_{p,N}^*(\mathbf{x})|\mathbf{x}\right),
$$

where $q_{p,N}^*(\mathbf{x})$ lies between $q_p(\mathbf{x})$ and $q_{p,N}(\mathbf{x})$. Hence, by standard arguments we can write

$$
|F\left(q_{p,N}(\mathbf{x})|\mathbf{x}\right) - F\left(q_p(\mathbf{x})|\mathbf{x}\right)| \leq 2 \sup_{a \leq y \leq b} |F_N(y|\mathbf{x}) - F(y|\mathbf{x})|,
$$

which implies that

$$
\sup_{\mathbf{x} \in \Omega_0} |q_{p,N}(\mathbf{x}) - q_p(\mathbf{x})| f\left(q_{p,N}^*(\mathbf{x})|\mathbf{x}\right) \leq 2 \sup_{\mathbf{x} \in \Omega_0} \sup_{a \leq y \leq b} |F_N(y|\mathbf{x}) - F(y|\mathbf{x})|.
$$

Thus, by Theorem 1 and because $f\left(q_p(\cdot)|\cdot\right)$ is uniformly lower-bounded away from zero, we assert that

$$
\sup_{\mathbf{x} \in \Omega_0} |q_{p,N}(\mathbf{x}) - q_p(\mathbf{x})| = O\left(\sqrt{\frac{\log N}{N h_K^d}} + h_K^2 + h_H^2\right) \text{ } \mathbb{P}\text{-p.s., as n} \to \infty.
$$

This ends the proof of Corollary 1. □

Bibliographie

BULINSKI A., SHASHKIN A., *Limit theorems for associated random fields and related systems.* Singapore World Scientific, 2007.

DOUKHAN P., NEUMANN M., Probability and moment inequalities for sums of weakly dependent random variables, with applications. *Stochastic Processes and their Applications* 117 (2007) : 878-903.

CHAUDHURI P., DOKSUM K., SAMAROV A., On average derivative quantile regression. *Ann. Statist.* 25 (1997) : 715–744.

ESARY J., PROSCHAN F., WALKUP D., Association of random variables with applications. *Ann. Math. Stat.* 38 (1967) : 1466-1476.

HONDA T., Nonparametric estimation of a conditional quantile for strong mixing processes. *Ann. Inst. Statist. Math.* 52 (2000) :459–470.

MEHRA K.L., RAO M.S., UPADRASTA S.P., A smooth conditional quantile estimator and related applications of conditional empirical processes. *J. Multivariate Anal.* 37 (1991) :151-179

OLIVEIRA P.E., *Asymptotics for Associated Random Variables.* Springer Verlag, 2012.

PRAKASA RAO B.L.S., *Associated Sequences, Demimartingales and Nonparametric Inference.* Probability and its Applications, Springer Basel AG, 2012

XIANG X., A kernel estimator of a conditional quantile, *J. Multivariate Anal.* 59 (1996) : 206-216.

YU K., JONES M.C., Local linear quantile regression. *J. Am. Stat. Assoc.* 93 (1998) : 228-238